6/19/20

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2003, 14:59:48 ; Search time 70 Seconds
(without alignments)
454.956 Million cell updates/sec

Title: US-09-887-784-4
Perfect score: 1274
Sequence: 1 MVSKGEELFTGVVPILVELD.....VLLGFVTAAGITLGMDELYK 239

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1274	100.0	239	23	AAE17518	Enhanced F64L-E222
2	1274	100.0	893	22	AAG65781	Amino acid sequenc
3	1274	100.0	1132	22	AAG65782	Amino acid sequenc
4	1266	99.4	239	23	AAE17517	Enhanced F64L jell
5	1263	99.1	239	21	AAB22882	Enhanced green flu
6	1263	99.1	239	21	AAY79584	EGFP signal domain
7	1263	99.1	239	21	AAY54349	Amino acid sequenc
8	1263	99.1	239	22	AAB31171	Amino acid sequenc
9	1263	99.1	239	22	AAB50804	Jellyfish GFP muta
10	1263	99.1	239	22	AAB85900	A. victoria green
11	1263	99.1	239	23	AAG66198	A. victoria green
12	1263	99.1	239	23	AAE14599	Aequorea victoria
13	1263	99.1	248	23	AAG68319	Jellyfish green fl
14	1263	99.1	259	23	AAU99804	Biomembrane permea
15	1263	99.1	265	20	AAW97451	Wild-type green fl
16	1263	99.1	268	23	AAU99803	Biomembrane permea
17	1263	99.1	270	23	AAU99802	Biomembrane permea
18	1263	99.1	272	23	AAU99800	Biomembrane permea
19	1263	99.1	273	23	AAU99801	Biomembrane permea
20	1263	99.1	280	23	AAU99807	Biomembrane permea
21	1263	99.1	281	21	AAB24252	EGFP-MODC422-461 f
22	1263	99.1	281	21	AAY50142	Green fluorescent
23	1263	99.1	281	23	AAU10888	EGFP-MODC422-461 f
24	1263	99.1	294	21	AAB22860	GFP-DEVD-annexin I
25	1263	99.1	294	21	AAY79638	Caspase-3 biosenso
26	1263	99.1	308	20	AAY42181	EGFP/DRM fusion pr
27	1263	99.1	323	21	AAY54359	GFP mutant EGFP li
28	1263	99.1	338	22	AAB86137	T. maritima HU/EGF
29	1263	99.1	359	20	AAY42179	EGFP/DRM fusion pr
30	1263	99.1	359	20	AAY42180	EGFP/DRM fusion pr
31	1263	99.1	364	20	AAW95356	IL-1 alpha propiec
32	1263	99.1	379	19	AAW85036	Alpha-actinin acti
33	1263	99.1	386	23	AAU96161	Human ARP/green fl
34	1263	99.1	391	20	AAY42178	EGFP/DRM fusion pr
35	1263	99.1	397	20	AAY42176	EGFP/DRM fusion pr
36	1263	99.1	403	20	AAY42177	EGFP/DRM fusion pr
37	1263	99.1	429	20	AAY42175	EGFP/DRM fusion pr
38	1263	99.1	442	19	AAW85035	Green fluorescent
39	1263	99.1	459	21	AAB22936	GFP-HSP27 fusion p
40	1263	99.1	544	19	AAW85028	CDK2-green fluores
41	1263	99.1	544	19	AAW85029	Green fluorescent
42	1263	99.1	552	21	AAY84883	A GFP-I-kappaB kin
43	1263	99.1	592	22	AAB85905	Amino acid sequenc
44	1263	99.1	604	19	AAW85016	Erk2-green floures
45	1263	99.1	604	22	AAB85904	Amino acid sequenc

ALIGNMENTS

RESULT 1
AAE17518
ID AAE17518 standard; Protein; 239 AA.
XX
AC AAE17518;
XX
DT 22-APR-2002 (first entry)
XX
DE Enhanced F64L-E222G jellyfish green fluorescent protein mutant.
XX
KW Jellyfish; green fluorescent protein; GFP; protein redistribution;
KW cellular function; genetic reporter; mutant; Stoke's shift; mutein.
XX
OS Aequorea victoria.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 65
FT /note= "Wild type Phe substituted with Leu; This
FT corresponds to position 64 in the wild type protein"
FT Misc-difference 223
FT /note= "Wild type Glu substituted with Gly; This
FT corresponds to position 222 in the wild type protein"
XX
PN WO200198338-A2.
XX
PD 27-DEC-2001.
XX
PF 18-JUN-2001; 2001WO-EP06848.
XX
PR 19-JUN-2000; 2000DK-0000953.
PR 20-JUN-2000; 2000US-212681P.
PR 10-MAY-2001; 2001DK-0000739.

PR 10-MAY-2001; 2001US-290170P.
XX
PA (BIOI-) BIOIMAGE AS.
XX
PI Bjorn SP, Pagliaro L, Thastrup O;
XX
DR WPI; 2002-098224/13.
DR N-PSDB; AAD28163.
XX
PT Novel fluorescent protein in in vitro assay for measuring protein
PT kinase activity or dephosphorylation activity, or for measuring protein
PT redistribution, has a green fluorescent protein with F64L and E222G
PT mutation -
XX
PS Claim 9; Page 37; 41pp; English.
XX
CC The invention relates to a fluorescent protein derived from green
CC fluorescent protein (GFP) or its analogue. The GFP containing mutations
CC at F64L and E222G has a bigger compared to other GFP's
CC making it very suitable for high throughput screening due to better
CC resolution. The fluorescent protein is useful in invitro assays for
CC measuring protein kinase activity or dephosphorylation activity, or for
CC measuring protein redistribution. The fluorescent protein is useful in
CC studying cellular functions in living cells; as protein tags in
CC transgenic animals, living and fixed cells; organelle tags, secretion
CC marker and genetic reporter. The fluorescent protein is also useful as
CC a cell or organelle integrity marker, a marker for changes in cell
CC morphology, as transfection marker, and as a marker to be used in
CC combination with fluorescence activated cell sorting (FACS). The novel
CC proteins can also be used as reporters to monitor live or dead biomass of
CC organisms, such as fungi. The fluorescent protein is also useful as
CC markers in transcriptional and translational fusions for performing
CC transposon vector mutagenesis and as a reporter for bacterial detection.
CC Transposons encoding the fluorescent protein are useful for screening
CC promoters and for tagging plasmids and chromosomes. The fluorescent
CC protein engineered into the genome of a phage is useful for designing
CC diagnostic tool. The present sequence is a DNA encoding enhanced
CC F64L-E222G jellyfish green fluorescent protein (GFP) mutant.
XX
SQ Sequence 239 AA;

Query Match 100.0%; Score 1274; DB 23; Length 239;
Best Local Similarity 100.0%; Pred. No. 2.6e-123;
Matches 239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPT 60

Qy         61 LVTTLSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVTTLSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL 120

Qy        121 VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA 180

Qy        181 DHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITLGMDELYK 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITLGMDELYK 239
```

RESULT 2
AAG65781
ID AAG65781 standard; Protein; 893 AA.
XX
AC AAG65781;
XX
DT 07-JAN-2002 (first entry)
XX
DE Amino acid sequence of HSPDE4A1-E222G fusion protein.
XX
KW PDE4; central nervous system; antiinflammatory; cytostatic; nootropic;
KW autoimmune; ischemic; osteopathic; GFP; green fluorescent protein;
KW fusion protein.
XX
OS Homo sapiens.
OS Aequorea victoria.
XX
PN WO200179526-A2.
XX
PD 25-OCT-2001.
XX
PF 11-APR-2001; 2001WO-DK00264.
XX
PR 17-APR-2000; 2000DK-0000651.
PR 29-MAY-2000; 2000DK-0000849.
XX
PA (BIOI-) BIOIMAGE AS.
XX
PI Terry BR, Scudder KM, Bjorn SP, Thastrup O, Almholt DC;
PI Praestegaard M;
XX
DR WPI; 2001-611727/70.
DR N-PSDB; AAI66852.
XX
PT Determining if a compound is a dislocator of PDE4 for identifying
PT compounds for treating CNS and inflammatory disease comprises
PT identifying compounds which remove PDE4 spots -
XX
PS Example 1; Page 156-160; 160pp; English.
XX
CC The invention relates to determining, if a compound, is a dislocator of
CC PDE4. The method comprises testing if the compound removes PDE4 spots,
CC which may optionally be induced by a Rolipram-like reference compound,
CC and testing if it inhibits the catalytic activity of the PDE4, where the
CC compound is a dislocator of PDE4, if it removes PDE spots and if it does
CC not inhibit the catalytic activity of PDE4. The method is useful for
CC identifying compounds useful for the treatment of diseases of the central
CC nervous system such as. depression and for the treatment of inflammatory
CC disease such as joint inflammation, Crohn's disease, inflammatory bowel
CC disease, respiratory diseases, chronic obstructive pulmonary disease
CC (COPD), including asthma, chronic bronchitis, pulmonary emphysema,
CC endotoxic shock, toxic shock syndrome, systemic lupus erthematosis,
CC psoriasis, bone resorption diseases, reperfusion injury, cancer and HIV
CC infection. The use of a reagent that can mimic or reverse the effect of
CC the compound with affinity for the catalytic site on intracellular
CC distribution of the PDE for the preparation of a medicament. The present
CC sequence represents the amino acid sequence of a HSPDE4A1-E222G fusion
CC protein.
XX
SQ Sequence 893 AA;

Query Match 100.0%; Score 1274; DB 22; Length 893;
Best Local Similarity 100.0%; Pred. No. 1.8e-122;
Matches 239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPT 714

Qy         61 LVTTLSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 LVTTLSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL 774

Qy        121 VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA 834

Qy        181 DHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITLGMDELYK 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        835 DHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITLGMDELYK 893
```

RESULT 3
AAG65782

```
ID   AAG65782 standard; Protein; 1132 AA.
XX
AC   AAG65782;
XX
DT   07-JAN-2002  (first entry)
XX
DE   Amino acid sequence of HSPDE4A4-E222G fusion protein.
XX
KW   PDE4; central nervous system; antiinflammatory; cytostatic; nootropic;
KW   autoimmune; ischemic; osteopathic; GFP; green fluorescent protein;
KW   fusion protein.
XX
OS   Homo sapiens.
OS   Aequorea victoria.
XX
PN   WO200179526-A2.
XX
PD   25-OCT-2001.
XX
PF   11-APR-2001; 2001WO-DK00264.
XX
PR   17-APR-2000; 2000DK-0000651.
PR   29-MAY-2000; 2000DK-0000849.
XX
PA   (BIOI-) BIOIMAGE AS.
XX
PI   Terry BR,  Scudder KM,  Bjorn SP,  Thastrup O,  Almholt DC;
PI   Praestegaard M;
XX
DR   WPI; 2001-611727/70.
DR   N-PSDB; AAI66853.
XX
PT   Determining if a compound is a dislocator of PDE4 for identifying
PT   compounds for treating CNS and inflammatory disease comprises
PT   identifying compounds which remove PDE4 spots -
XX
PS   Example 1; Page 162-167; 160pp; English.
XX
CC   The invention relates to determining, if a compound, is a dislocator of
CC   PDE4. The method comprises testing if the compound removes PDE4 spots,
CC   which may optionally be induced by a Rolipram-like reference compound,
CC   and testing if it inhibits the catalytic activity of the PDE4, where the
CC   compound is a dislocator of PDE4, if it removes PDE spots and if it does
CC   not inhibit the catalytic activity of PDE4. The method is useful for
CC   identifying compounds useful for the treatment of diseases of the central
CC   nervous system such as. depression and for the treatment of inflammatory
CC   disease such as joint inflammation, Crohn's disease, inflammatory bowel
CC   disease, respiratory diseases, chronic obstructive pulmonary disease
CC   (COPD), including asthma, chronic bronchitis, pulmonary emphysema,
CC   endotoxic shock, toxic shock syndrome, systemic lupus erthematosis,
CC   psoriasis, bone resorption diseases, reperfusion injury, cancer and HIV
CC   infection. The use of a reagent that can mimic or reverse the effect of
CC   the compound with affinity for the catalytic site on intracellular
CC   distribution of the PDE for the preparation of a medicament. The present
CC   sequence represents the amino acid sequence of a HSPDE4A4-E222G fusion
CC   protein.
XX
SQ   Sequence   1132 AA;

  Query Match             100.0%;  Score 1274;  DB 22;  Length 1132;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-122;
  Matches  239;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        894 MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPT 953

Qy         61 LVTTLSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        954 LVTTLSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL 1013

Qy        121 VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       1014 VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA 1073

Qy        181 DHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITLGMDELYK 239
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1074 DHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITLGMDELYK 1132


RESULT 4
AAE17517
ID   AAE17517 standard; Protein; 239 AA.
XX
AC   AAE17517;
XX
DT   22-APR-2002  (first entry)
XX
DE   Enhanced F64L jellyfish green fluorescent protein mutant.
XX
KW   Jellyfish; green fluorescent protein; GFP; protein redistribution;
KW   cellular function; genetic reporter; mutant; Stoke's shift; mutein.
XX
OS   Aequorea victoria.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 65
FT                   /note= "Wild type Phe substituted with Leu; This
FT                   corresponds to position 64 in the wild type protein"
XX
PN   WO200198338-A2.
XX
PD   27-DEC-2001.
XX
PF   18-JUN-2001; 2001WO-EP06848.
XX
PR   19-JUN-2000; 2000DK-0000953.
PR   20-JUN-2000; 2000US-212681P.
PR   10-MAY-2001; 2001DK-0000739.
PR   10-MAY-2001; 2001US-290170P.
XX
PA   (BIOI-) BIOIMAGE AS.
XX
PI   Bjorn SP,  Pagliaro L,  Thastrup O;
XX
DR   WPI; 2002-098224/13.
DR   N-PSDB; AAD28162.
XX
PT   Novel fluorescent protein in in vitro assay for measuring protein
PT   kinase activity or dephosphorylation activity, or for measuring protein
PT   redistribution, has a green fluorescent protein with F64L and E222G
PT   mutation  -
XX
PS   Example 1; Page 35; 41pp; English.
XX
CC   The invention relates to a fluorescent protein derived from green
CC   fluorescent protein (GFP) or its analogue. The GFP containing mutations
CC   at F64L and E222G has a bigger compared to other GFP's
CC   making it very suitable for high throughput screening due to better
CC   resolution. The fluorescent protein is useful in invitro assays for
CC   measuring protein kinase activity or dephosphorylation activity, or for
CC   measuring protein redistribution. The fluorescent protein is useful in
CC   studying cellular functions in living cells; as protein tags in
CC   transgenic animals, living and fixed cells; organelle tags, secretion
CC   marker and genetic reporter. The fluorescent protein is also useful as
CC   a cell or organelle integrity marker, a marker for changes in cell
CC   morphology, as transfection marker, and as a marker to be used in
CC   combination with fluorescence activated cell sorting (FACS). The novel
CC   proteins can also be used as reporters to monitor live or dead biomass of
CC   organisms, such as fungi. The fluorescent protein is also useful as
CC   markers in transcriptional and translational fusions for performing
CC   transposon vector mutagenesis and as a reporter for bacterial detection.
CC   Transposons encoding the fluorescent protein are useful for screening
CC   promoters and for tagging plasmids and chromosomes. The fluorescent
CC   protein engineered into the genome of a phage is useful for designing
```

CC diagnostic tool. The present sequence is enhanced F64L jellyfish green
CC fluorescent protein (GFP) mutant.
XX
SQ Sequence 239 AA;

Query Match 99.4%; Score 1266; DB 23; Length 239;
Best Local Similarity 99.6%; Pred. No. 1.8e-122;
Matches 238; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPT 60

Qy         61 LVTTLSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVTTLSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL 120

Qy        121 VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA 180

Qy        181 DHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITLGMDELYK 239
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        181 DHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK 239
```

RESULT 5
AAB22882
ID AAB22882 standard; Protein; 239 AA.
XX
AC AAB22882;
XX
DT 10-JAN-2001 (first entry)
XX
DE Enhanced green fluorescent protein (EGFP), SEQ ID NO:46.
XX
KW Biodetector protein; fusion protein; recognition site;
KW cellular targetting sequence; cellular localisation; fluorescent protein;
KW protease activity detection; toxin detection; cellular stress detection;
KW drug discovery; cell based screening.
XX
OS Aequorea victoria.
OS Synthetic.
XX
PN WO200050872-A2.
XX
PD 31-AUG-2000.
XX
PF 25-FEB-2000; 2000WO-US04794.
XX
PR 26-FEB-1999; 99US-0122152.
PR 08-MAR-1999; 99US-0123399.
PR 12-JUL-1999; 99US-0352171.
XX
PA (CELL-) CELLOMICS INC.
XX
PI Giuliano KA, Kapur R;
XX
DR WPI; 2000-594086/56.
DR N-PSDB; AAA93373.
XX
PT Automated cell-based characterization of toxin by contacting cells
PT containing luminescent reporter molecules with test substance and
PT analyzing optically -
XX
PS Example 11; Fig 29A; 336pp; English.
XX
CC The invention relates to systems, methods and reagents for cell-based
CC screening or detection of compounds which affect particular biological
CC functions. The methods of the invention utilise fluorescent biodetector
CC molecules which, when acted on by a compound of interest, cause an
CC alteration in the cellular distribution of at least the fluorescent
CC moiety. In one embodiment, the biosensors comprise heat shock proteins
CC (HSPs) fused to a fluorescent protein (e.g., jellyfish green fluorescent
CC protein (GFP), or derivatives thereof). Such biosensors are located in
CC the cytoplasm, but on stress activation translocate to the nucleus. In
CC another embodiment biodetector proteins can be used to detect protease
CC activity. Such protease biodetector fusion proteins comprise one or more
CC fluorescent proteins; a recognition signal which is cleaved by the
CC protease; and at least one cellular localisation signal. The latter two
CC components may be components of a single protein which is acted upon by
CC the protease, or may be from heterologous sources. Due to the
CC localisation signal, the biodetector protein is localised to a
CC particular region of the cell. Once acted on by the protease of interest,
CC the fluorescent protein is cleaved from the localisation sequence, and
CC is free to migrate to other locations within the cell. The presence of a
CC second localisation signal attached to the fluorescent protein enables
CC the fluorescent protein to be directed to a different cellular
CC compartment after cleavage of the protease recognition sequence. The
CC change in distribution of the fluorescent protein can be detected using
CC imaging methods with a high degree of spatial resolution. The methods
CC and biosensors of the invention can be used to investigate a wide range
CC of cellular activities and to screen compounds which modulate these
CC activities. Biosensors containing a recognition site for caspase, for
CC example, may be used for the screening of compounds which modulate
CC apoptosis, while biosensors containing other protease recognition sites
CC may be used for the detection of proteolytic toxins (such as anthrax
CC lethal factor). The method provides improved target validation and
CC candidate compound optimisation by combining many cell screening formats
CC with fluorescence-based molecular reagents and computer-based feature
CC extraction, data analysis and automation, resulting in increased
CC quantity and speed of data collection and faster evaluation of drug
CC candidates. Sequences AAB22881-B22885 represent fluorescent proteins
CC which may used as components of biosensor fusion proteins of the
CC invention.
XX
SQ Sequence 239 AA;

Query Match 99.1%; Score 1263; DB 21; Length 239;
Best Local Similarity 99.2%; Pred. No. 3.6e-122;
Matches 237; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPT 60

Qy         61 LVTTLSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL 120
              |||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL 120

Qy        121 VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA 180

Qy        181 DHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITLGMDELYK 239
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        181 DHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK 239
```

RESULT 6
AAY79584
ID AAY79584 standard; Peptide; 239 AA.
XX
AC AAY79584;
XX
DT 29-AUG-2000 (first entry)
XX
DE EGFP signal domain.
XX
KW Protease; biosensor; EGFP; signal peptide; cell screening;
KW assay; analysis; drug discovery.
XX
OS Unidentified.
XX
PN WO200026408-A2.
XX

PD 11-MAY-2000.
XX
PF 29-OCT-1999; 99WO-US25431.
XX
PR 30-OCT-1998; 98US-0106308.
PR 26-MAY-1999; 99US-0136078.
XX
PA (CELL-) CELLOMICS INC.
XX
PI Guiliano KA, Bright G, Olson K, Burroughs-Tencza S;
XX
DR WPI; 2000-365644/31.
DR N-PSDB; AAA27573.
XX
PT Recombinant nucleic acid encoding a protease biosensor useful for
PT fluorescence based cell and molecular biochemical assays for drug
PT discovery comprising three operably linked nucleic acid sequences -
XX
PS Claim 14; Fig 29A; 218pp; English.
XX
CC The present sequence is that of the EGFP signal domain, which can
CC be included in novel recombinant protease biosensors (PBs) of the
CC invention. The PBs (see AAY79638-54) comprise: a first domain (see
CC AAY79579-87) comprising at least 1 detectable polypeptide signal
CC such as the present sequence; a second domain (see AAY79588-622)
CC comprising at least 1 protease recognition site; and a third domain
CC (see AAY79623-37) comprising at least 1 reactant target sequence. A
CC recombinant nucleic acid (see AAA27627-43) encoding the PB, an
CC expression vector, and a genetically engineered host cell are also
CC claimed. A claimed method for identifying compounds that modify
CC protease activity in a cell involves contacting a host cell that
CC possesses the recombinant PB with a test compound, and determining
CC the PB distribution in the host cell, where changes in the
CC distribution of the PB are correlated with modification of protease
CC activity by the test compound. Claimed kits for identifying
CC compounds that modify protease activity in a host cell include the
CC recombinant nucleic acid, or the recombinant PB, or the vector, or
CC the host cell. The PB is useful in high content screens to detect
CC in vivo activation of enzymatic activity, and to identify specific
CC activity based on cleavage of a known recognition motif.
XX
SQ Sequence 239 AA;

Query Match 99.1%; Score 1263; DB 21; Length 239;
Best Local Similarity 99.2%; Pred. No. 3.6e-122;
Matches 237; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPT 60

Qy         61 LVTTLSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL 120
              |||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL 120

Qy        121 VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA 180

Qy        181 DHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITLGMDELYK 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK 239
```

RESULT 7
AAY54349
ID AAY54349 standard; Protein; 239 AA.
XX
AC AAY54349;
XX
DT 06-APR-2000 (first entry)
XX
DE Amino acid sequence of the mutant green fluorescent protein EGFP.
XX
KW Fluorescent protein; green fluorescent protein; emission intensity;
KW fluorescence; pH detection; pH sensor; EGFP.
XX
OS Synthetic.
OS Aequorea victoria.
XX
FH Key Location/Qualifiers
FT Misc-difference 65
FT /note= "wild type Phe substituted with Leu"
FT Misc-difference 66
FT /note= "wild type Ser substituted with Thr"
FT Misc-difference 232
FT /note= "wild type His substituted with Leu"
XX
PN WO9964592-A2.
XX
PD 16-DEC-1999.
XX
PF 08-JUN-1999; 99WO-US12850.
XX
PR 09-JUN-1998; 98US-0094359.
PR 13-OCT-1998; 98US-0172063.
XX
PA (REGC) UNIV CALIFORNIA.
PA (UYOR-) UNIV OREGON STATE.
XX
PI Tsien RY, Llopis J, Wachter RM;
XX
DR WPI; 2000-116540/10.
DR N-PSDB; AAZ45642.
XX
PT New functional engineered green fluorescent proteins, used for
PT measuring the pH in biological samples and cells -
XX
PS Disclosure; Page 9; 89pp; English.
XX
CC The present sequence represents a functional engineered fluorescent
CC protein based on the Aequorea green fluorescent protein (GFP). The
CC emission intensity changes as pH varies between 5 and 10 of the
CC present protein are novel. The functional engineered fluorescent
CC proteins show reversible changes in fluorescence over physiological
CC pH ranges. They can be used for determining the pH of samples and
CC cells. The polynucleotides can also be used to produce transgenic
CC animals. The fluorescent protein pH sensors can be delivered to
CC cells in the form of polynucleotides encoding the protein sensor
CC fused to a targeting signal. The targeting signal directs the
CC expression of the protein sensors to restricted cell locations. This
CC makes it possible to measure the pH of a precisely defined cellular
CC region or organelle.
XX
SQ Sequence 239 AA;

Query Match 99.1%; Score 1263; DB 21; Length 239;
Best Local Similarity 99.2%; Pred. No. 3.6e-122;
Matches 237; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPT 60

Qy         61 LVTTLSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL 120
              |||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL 120

Qy        121 VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA 180

Qy        181 DHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITLGMDELYK 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK 239
```

```
RESULT 8
AAB31171
ID   AAB31171 standard; Protein; 239 AA.
XX
AC   AAB31171;
XX
DT   02-APR-2001  (first entry)
XX
DE   Amino acid sequence of a green fluorescent protein (GFP).
XX
KW   Growth rate; death rate; reporter gene; luminescent protein;
KW   fluorescent product; luciferase; green fluorescent protein; GFP.
XX
OS   Aequorea victoria.
XX
PN   WO200075367-A1.
XX
PD   14-DEC-2000.
XX
PF   07-JUN-2000; 2000WO-FI00507.
XX
PR   07-JUN-1999;   99FI-0001296.
XX
PA   (LILI/) LILIUS E.
PA   (VIRT/) VIRTA M.
XX
PI   Lilius E,  Virta M;
XX
DR   WPI; 2001-061737/07.
DR   N-PSDB; AAC86954.
XX
PT   Assessing growth and death rates of a micro-organism in a desired
PT   environment, by introducing 2 reporter genes encoding luminescent and
PT   fluorescent products and detecting luminescent fluorescence  -
XX
PS   Disclosure; Page 27; 32pp; English.
XX
CC   The specification describes a method for assessing the growth rate
CC   and death rate of a micro-organism within a predetermined time period
CC   in a desired environment. The method comprises introducing at least
CC   two reporter genes encoding luminescent and/or fluorescent products
CC   into the micro-organisms, incubating the micro-organism within the
CC   desired environment, and detecting luminescence and/or fluorescence
CC   after a predetermined time period. Use of two different markers within
CC   a micro-organism enables the differentiation between growth and death
CC   rates. The method is used to assess the growth rate and death rate of
CC   a micro-organism within a predetermined time period in a desired
CC   environment. The present sequence represents a green fluorescent
CC   protein (GFP), and is encoded by a plasmid which encodes luminescent
CC   and fluorescent proteins, and is used in the method of the invention.
XX
SQ   Sequence   239 AA;

  Query Match            99.1%;  Score 1263;  DB 22;  Length 239;
  Best Local Similarity  99.2%;  Pred. No. 3.6e-122;
  Matches  237;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPT 60

Qy         61 LVTTLSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL 120
              |||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL 120

Qy        121 VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA 180

Qy        181 DHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITLGMDELYK 239
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        181 DHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK 239
```

```
RESULT 9
AAB50804
ID   AAB50804 standard; Protein; 239 AA.
XX
AC   AAB50804;
XX
DT   14-MAR-2001  (first entry)
XX
DE   Jellyfish GFP mutant EGFP.
XX
KW   Aequorea victoria; jellyfish; fluorescent protein indicator;
KW   green fluorescent protein; GFP; linker moiety; sensor;
KW   calmodulin-binding domain; mutant; mutein.
XX
OS   Aequorea victoria.
XX
PN   WO200071565-A2.
XX
PD   30-NOV-2000.
XX
PF   17-MAY-2000; 2000WO-US13684.
XX
PR   21-MAY-1999;   99US-0316919.
PR   21-MAY-1999;   99US-0316920.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Tsien RY,  Baird GA;
XX
DR   WPI; 2001-032017/04.
DR   N-PSDB; AAC90488.
XX
PT   Novel fluorescent proteins comprising a sensor protein inserted into
PT   them, useful for measuring the response of a sensor biological,
PT   chemical, electrical or physiological parameter in vivo or in vitro  -
XX
PS   Disclosure; Page 24; 94pp; English.
XX
CC   The present sequence is a fluorescnet protein used in the construction
CC   of a fluorescent protein indicator. The indicator comprises a sensor
CC   polypeptide that is responsive to a chemical, biological, electrical
CC   or physiological parameter, and a fluorescence protein functional group.
CC   The sensor polypeptide is operatively inserted into the fluorescent
CC   moiety. The fluorescent indicator is useful for detecting the presence of
CC   a response inducing member in a sample. The method involves contacting
CC   the sample with the indicator and detecting a change in fluorescence, in
CC   which a change is indicative of the effect of the parameter on the sensor
CC   polypeptide. The novel fluorescent proteins are advantageous due to their
CC   reduced size as compared to the FRET (fluorescence resonance energy
CC   transfer)-based sensors.
XX
SQ   Sequence   239 AA;

  Query Match            99.1%;  Score 1263;  DB 22;  Length 239;
  Best Local Similarity  99.2%;  Pred. No. 3.6e-122;
  Matches  237;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPT 60

Qy         61 LVTTLSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL 120
              |||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL 120

Qy        121 VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA 180

Qy        181 DHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITLGMDELYK 239
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        181 DHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK 239
```

```
Db        181 DHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK 239


RESULT 10
AAB85900
ID   AAB85900 standard; Protein; 239 AA.
XX
AC   AAB85900;
XX
DT   30-NOV-2001  (first entry)
XX
DE   A. victoria green fluorescent protein (GFP) and linker sequence.
XX
KW   Melanin concentrating hormone receptor; MCHR; MCH; chimeric; fusion;
KW   fluorescent polypeptide; orexigenic; anabolic; food intake; GFP;
KW   green fluorescent protein.
XX
OS   Synthetic.
OS   Aequorea victoria.
XX
PN   WO200168706-A1.
XX
PD   20-SEP-2001.
XX
PF   14-MAR-2001; 2001WO-US08071.
XX
PR   15-MAR-2000; 2000US-0189698.
XX
PA   (MERI ) MERCK & CO INC.
XX
PI   Marsh DJ;
XX
DR   WPI; 2001-565791/63.
DR   N-PSDB; AAH47304.
XX
PT   Fusion proteins comprising melanin concentrating hormone receptor
PT   peptides and fluorescent proteins, useful for identifying appetite
PT   stimulants -
XX
PS   Claim 2; Page 14; 71pp; English.
XX
CC   The invention provides melanin concentrating hormone (MCH) receptor
CC   (MCHR) chimeric and fusion proteins. The MCHR chimeric proteins comprise
CC   MCHR polypeptide regions from different species. The MCHR fusion protein
CC   comprise MCHR polypeptide region and a fluorescent polypeptide region
CC   joined directly, or via a linker, to the carboxy side of the MCHR
CC   polypeptide region. The MCHR fusion proteins can be expressed by standard
CC   recombinant methodology. MCH action promotes feeding (orexigenic) and up
CC   regulation of MCH activity stimulates food intake. The present sequence
CC   represents a A. victoria green fluorescent protein (GFP) and a linker
CC   sequence.
XX
SQ   Sequence   239 AA;

  Query Match             99.1%;  Score 1263;  DB 22;  Length 239;
  Best Local Similarity   99.2%;  Pred. No. 3.6e-122;
  Matches  237;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPT 60

Qy         61 LVTTLSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL 120
              |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL 120

Qy        121 VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA 180

Qy        181 DHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITLGMDELYK 239
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        181 DHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK 239


RESULT 11
AAG66198
ID   AAG66198 standard; Protein; 239 AA.
XX
AC   AAG66198;
XX
DT   17-JUN-2002  (first entry)
XX
DE   A. victoria green fluorescent protein (EGFP).
XX
KW   Cyan-green fluorescent protein; fluorescence; recombinant; GFP;
KW   green fluorescent protein; EGFP.
XX
OS   Aequorea victoria.
XX
PN   JP2002045189-A.
XX
PD   12-FEB-2002.
XX
PF   04-AUG-2000; 2000JP-0237165.
XX
PR   04-AUG-2000; 2000JP-0237165.
XX
PA   (RIKA ) RIKAGAKU KENKYUSHO.
XX
DR   WPI; 2002-299190/34.
DR   N-PSDB; ABL40628.
XX
PT   A gene encoding cyan-green fluorescent protein  -
XX
PS   Examples; Page 14; 20pp; Japanese.
XX
CC   The invention relates to a gene encoding proteins having cyan-green
CC   fluorescence characteristic and having a function of showing stable
CC   fluorescence characteristic in acid region. A method for the preparation
CC   of a cyan-green fluorescent protein is provided which involves a
CC   transformant transformed by a recombinant vector comprising the gene,
CC   where the transformant is cultured and the protein is collected from the
CC   culture. The present sequence represents the A. victoria green
CC   fluorescent protein (EGFP).
XX
SQ   Sequence   239 AA;

  Query Match             99.1%;  Score 1263;  DB 23;  Length 239;
  Best Local Similarity   99.2%;  Pred. No. 3.6e-122;
  Matches  237;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPT 60

Qy         61 LVTTLSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL 120
              |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL 120

Qy        121 VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA 180

Qy        181 DHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITLGMDELYK 239
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        181 DHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK 239


RESULT 12
AAE14599
ID   AAE14599 standard; Protein; 239 AA.
XX
AC   AAE14599;
XX
DT   31-MAY-2002  (first entry)
```

```
XX
DE   Aequorea victoria enhanced green fluorescent protein.
XX
KW   Mutagenesis; enhanced green fluorescent protein; EGFP; mutant; mutein.
XX
OS   Aequorea victoria.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1..3
FT                   /note= "Wild-type GFP Met-Ser are replaced with
FT                   Met-Val-Ser"
FT   Misc-difference 65
FT                   /note= "GFP Phe64 is replaced by Leu"
FT   Misc-difference 66
FT                   /note= "GFP Ser65 is replaced by Thr"
XX
PN   EP1178109-A1.
XX
PD   06-FEB-2002.
XX
PF   03-AUG-2001; 2001EP-0306650.
XX
PR   04-AUG-2000; 2000JP-0237166.
XX
PA   (RIKE ) RIKEN KK.
XX
PI   Miyawaki A,  Sawano A;
XX
DR   WPI; 2002-208112/27.
DR   N-PSDB; AAD27910.
XX
PT   Method for mutagenesis, e.g. for introducing certain or random
PT   mutations at certain sites of the nucleotide sequence, comprises
PT   synthesizing a mutated strand and a complementary strand by use of
PT   megaprimers  -
XX
PS   Example 1; Page 13-14; 31pp; English.
XX
CC   The invention relates to a method for mutagenesis that comprises
CC   synthesising a mutated strand and a complementary strand by use of
CC   megaprimers. The method basically comprises a DNA synthesis in which
CC   one or more primers that have a nucleotide sequence containing at least
CC   one mutation and a phosphorylated 5'-terminus are annealed to a template
CC   DNA and then subjected to an elongation reaction using a thermostable
CC   high-fidelity DNA polymerase, after which the phosphorylated 5'-terminus
CC   and the elongated terminus are ligated by means of a thermostable DNA
CC   ligase to synthesise a circular DNA containing the primers; a digestion
CC   in which the step of DNA synthesis is repeated several times to amplify
CC   the DNA containing the primers and then, at least DNAs other than the
CC   amplified circular DNA are digested into several fragments; and
CC   a double-stranded DNA synthesis in which, with the several
CC   fragments obtained in the step of digestion as megaprimers, the
CC   megaprimers are annealed to the circular DNA synthesised above,
CC   followed by an elongation reaction performed using the thermostable
CC   high-fidelity DNA polymerase. The method is useful for mutagenesis,
CC   particularly for introducing certain mutations at certain sites of a
CC   nucleotide sequence, or for introducing random mutations at certain sites
CC   of the nucleotide sequence. The present method is simple, speedy,
CC   economical and widely applicable. The present sequence is
CC   Aequorea victoria enhanced green fluorescent protein (EGFP)
CC   used for mutagenesis in an exemplification of the invention.
CC   The EGFP is derived by altering the green fluorescent protein (GFP)
CC   sequence of Aequorea victoria.
XX
SQ   Sequence   239 AA;

  Query Match            99.1%;  Score 1263;  DB 23;  Length 239;
  Best Local Similarity  99.2%;  Pred. No. 3.6e-122;
  Matches  237;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db          1 MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPT 60

Qy         61 LVTTLSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL 120
              |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db         61 LVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL 120

Qy        121 VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        121 VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA 180

Qy        181 DHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITLGMDELYK 239
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        181 DHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK 239


RESULT 13
AAG68319
ID   AAG68319 standard; Protein; 248 AA.
XX
AC   AAG68319;
XX
DT   21-FEB-2002  (first entry)
XX
DE   Jellyfish green fluorescent protein (GFP) SEQ ID NO:8.
XX
KW   Human; beta-amyloid; cyclin-dependent kinase inhibitor; nerve cell;
KW   amyloid precursor protein; APP; jellyfish; green fluorescent protein.
XX
OS   Unidentified.
XX
PN   WO200182967-A1.
XX
PD   08-NOV-2001.
XX
PF   25-APR-2001; 2001WO-JP03555.
XX
PR   28-APR-2000; 2000JP-0131037.
XX
PA   (YAMA ) YAMANOUCHI PHARM CO LTD.
PA   (SUZU/) SUZUKI T.
XX
PI   Suzuki T,  Watanabe T,  Kawabata S,  Hachiya S;
XX
DR   WPI; 2002-026209/03.
DR   N-PSDB; ABA04089.
XX
PT   Medicinal compositions for the treatment of dementia and Alzheimer's
PT   disease, comprise compounds that suppress beta amyloid production  -
XX
PS   Example 1; Page 48-50; 62pp; Japanese.
XX
CC   The present invention describes medicinal compositions (I) inhibiting
CC   beta-amyloid production comprising an active component a substance that
CC   inhibits the activity of cyclin-dependent kinase (CDK). Also described
CC   are: (1) a method for screening compounds for their ability to inhibit
CC   the production of beta-amyloid by contacting with beta-amyloid producing
CC   cells; and (2) screening kits. (I) have nootropic and neuroprotective
CC   activities. (I) suppress the phosphorylation of amyloid precursor protein
CC   (APP) which is an essential step in the production of beta-amyloid. (I)
CC   can be used in the treatment and prevention of neurodegenerative diseases
CC   such as dementia and Alzheimer's disease. The present sequence represents
CC   a jellyfish green fluorescent protein (GFP), which is used in the
CC   exemplification of the present invention.
XX
SQ   Sequence   248 AA;

  Query Match            99.1%;  Score 1263;  DB 23;  Length 248;
  Best Local Similarity  99.2%;  Pred. No. 3.8e-122;
  Matches  237;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db          1 MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPT 60
```

```
Qy         61 LVTTLSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL 120
              |||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL 120

Qy        121 VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA 180

Qy        181 DHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITLGMDELYK 239
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        181 DHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK 239
```

RESULT 14
AAU99804
ID AAU99804 standard; Protein; 259 AA.
XX
AC AAU99804;
XX
DT 07-OCT-2002 (first entry)
XX
DE Biomembrane permeable compound associated EGFP-Histidine tag protein.
XX
KW Biomembrane permeating signal sequence; nucleus-transfer signal;
KW postsynapse transfer signal sequence; biomembrane permeable compound;
KW PCR; primer; ss; enhanced green fluorescent protein; EGFP; His tag;
KW histidine tag.
XX
OS Synthetic.
XX
PN JP2002153288-A.
XX
PD 28-MAY-2002.
XX
PF 24-NOV-2000; 2000JP-0358442.
XX
PR 24-NOV-2000; 2000JP-0358442.
XX
PA (MATS/) MATSUI H.
PA (MATS/) MATSUSHITA M.
XX
DR WPI; 2002-552745/59.
XX
PT Compound for introducing a substance to a specific site in a cell, a
PT PAK inhibitor, a transcription inhibitor, a vector -
XX
PS Example 1; Page 16; 25pp; Japanese.
XX
CC The invention describes a compound containing a biomembrane permeating
CC signal sequence and a selectively introduced signal sequence to a
CC specific site in a cell and which can be localised in the specific site
CC in the cell. The biomembrane permeating signal sequence consists of
CC 9-13 arginine residues and the selectively introduced signal sequence
CC to a specific site in a cell is a nucleus-transfer or postsynapse
CC transfer signal sequence. The compound is used for localising a peptide
CC acting as a drug in a cell nucleus. This is the amino acid sequence of
CC an enhanced green fluorescent protein (EGFP) protein fused to a
CC histidine tag used in the development of a biomembrane permeable
CC compound.
XX
SQ Sequence 259 AA;

Query Match 99.1%; Score 1263; DB 23; Length 259;
Best Local Similarity 99.2%; Pred. No. 4.1e-122;
Matches 237; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPT 60

Qy         61 LVTTLSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL 120
              |||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         61 LVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL 120

Qy        121 VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA 180

Qy        181 DHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITLGMDELYK 239
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        181 DHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK 239
```

RESULT 15
AAW97451
ID AAW97451 standard; Protein; 265 AA.
XX
AC AAW97451;
XX
DT 19-MAY-1999 (first entry)
XX
DE Wild-type green fluorescent protein (GFP).
XX
KW Green fluorescent protein; GFP; fluorescent energy transfer;
KW G protein-coupled receptor; G protein; receptor.
XX
OS Aequorea victoria.
XX
PN WO9855873-A2.
XX
PD 10-DEC-1998.
XX
PF 04-JUN-1998; 98WO-FR01136.
XX
PR 05-JUN-1997; 97FR-0006977.
XX
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Alix P, Galzi JL;
XX
DR WPI; 1999-142415/12.
DR N-PSDB; AAX16086.
XX
PT Detecting non-covalent interactions between target protein and
PT ligand - using green fluorescent protein as energy transfer label
PT for reactants, used to, e.g. identify potential therapeutics binding
PT to G protein-coupled receptors
XX
PS Disclosure; Fig 1; 103pp; French.
XX
CC The present sequence represents wild-type green fluorescent protein
CC (GFP). The specification describes the use of GFP, or its fluorescent
CC variants and derivatives, for detecting and quantifying non-covalent
CC interactions between a target protein, genetically labelled by GFP,
CC and a ligand labelled with a group that can absorb light emitted from
CC GFP or is a fluorescent substance. The method is based on fluorescent
CC energy transfer between GFP and the group, with the fluorescent
CC substance being excited by light emitted from GFP or emitting at
CC the GFP excitation wavelength. The labelled reagents are especially
CC used to assess interaction between a G protein-coupled receptor and a
CC G protein, particularly to identify agents interacting reversibly at
CC the receptor, i.e. potential therapeutic agonists and antagonists.
XX
SQ Sequence 265 AA;

Query Match 99.1%; Score 1263; DB 20; Length 265;
Best Local Similarity 99.2%; Pred. No. 4.2e-122;
Matches 237; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPT 60

Qy         61 LVTTLSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL 120
              |||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        61 LVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTL 120

Qy       121 VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 VNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLA 180

Qy       181 DHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLGFVTAAGITLGMDELYK 239
            |||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db       181 DHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGMDELYK 239


Search completed: June  3, 2003, 15:06:36
Job time : 71 secs
```